CLAIMS

1   1.   A method for detecting and resolving a partition condition in a cluster of

2   computers in a networked environment, the method comprising:

3   creating a scratch pad area accessible by the cluster of computers;

4   dividing the scratch pad into a plurality of slots, each slot associated with a

5   plurality of nodes within the cluster of computers;

6   recording in the plurality of slots, a generation number and a list of known nodes

7   by each one of the plurality of nodes, wherein an identifier is written in the list for each

8   node that is known to a writing node and wherein the generation number and the list of

9   known nodes is recorded when a change of membership occurs in the cluster of

10   computers;

11   comparing each slot of the plurality of slots to ensure the generation number and

12   the list of known nodes matches in each slot of the plurality of slots; and

13   resolving the partition condition by creating a list of surviving nodes and re-

14   allocating appropriate resources to each of the surviving nodes.

1   2.   The method of claim 1 wherein the creating the list of surviving nodes includes

2   listing a first set of nodes determined by comparing each slot of the plurality of slots.

1   3.   The method of claim 2 wherein the comparing each slot includes finding a list

2   with a master node to create the list of surviving nodes and shutting down each node not

3   on the list with the master node.

1   4.   The method of claim 2 wherein the comparing each slot includes finding a list

2   with a lowest node rank to create the list of surviving nodes and shutting down each node

3   not on the list with the lowest node rank.

1   5.   The method of claim 2 wherein the comparing each slot includes finding a list

2   with a largest node to create the list of surviving nodes and shutting down each node not

3   on the list with the largest node.

1   6.   The method of claim 2 wherein the comparing each slot includes finding a list

2   with a maximum number of nodes to create the list of surviving nodes and shutting down

3   each node not on the list with the maximum number of nodes.

1   7.   The method of claim 2 further including shutting down each node not on the list

2   of surviving nodes by requiring each node not on the list of surviving nodes to write a

3   special message in a respective slot for that node and then shut down immediately.

1   8.   The method of claim 7 further including sending the list of surviving nodes to the

2   each node on the list of surviving nodes along with a new generation number.

1   9.   The method of claim 7 further including requiring each node not on the list of

2   surviving nodes to re-register with the cluster of computers.

1   10.   A computer program for detecting and resolving a partition condition in a cluster

2   of computers in a networked environment, the computer program comprising:

3   instructions for creating a scratch pad area accessible by the cluster of computers;

4   instructions for dividing the scratch pad into a plurality of slots, each slot

5   associated with a plurality of nodes within the cluster of computers;

6   instructions for recording in the plurality of slots, a generation number and a list

7   of known nodes by each one of the plurality of nodes, wherein an identifier is written in

8   the list for each node that is known to a writing node and wherein the generation number

9   and the list of known nodes is recorded when a change of membership occurs in the

10   cluster of computers;

11   instructions for comparing each slot of the plurality of slots to ensure the

12   generation number and the list of known nodes matches in each slot of the plurality of

13   slots; and

14   instructions for resolving the partition condition by creating a list of surviving

15   nodes and re-allocating appropriate resources to each of the surviving nodes.

1   11.   The computer program of claim 10 wherein the instructions for creating the list of

2   surviving nodes includes instructions for listing a first set of nodes determined by

3   comparing each slot of the plurality of slots

1   12.   The computer program of claim 11 wherein the instructions for comparing each

2   slot includes instructions for finding a list with a master node to create the list of

3   surviving nodes and shutting down each node not on the list with the master node.

1   13.   The computer program of claim 11 wherein the instructions for comparing each

2   slot includes instructions for finding a list with a lowest node rank to create the list of

3   surviving nodes and shutting down each node not on the list with the lowest node rank.

1   14.   The computer program of claim 11 wherein the instructions for comparing each

2   slot includes instructions for finding a list with a largest node to create the list of

3   surviving nodes and shutting down each node not on the list with the largest node.

1   15.   The computer program of claim 11 wherein the instructions for comparing each

2   slot includes instructions for finding a list with a maximum number of nodes to create the

3   list of surviving nodes and shutting down each node not on the list with the maximum

4   number of nodes.

1   16.   The computer program of claim 15 further including instructions for sending the

2   list of surviving nodes to the each node on the list of surviving nodes along with a new

3   generation number.

1   17.   The computer program of 16 further including requiring each node not on the list

2   of surviving nodes to re-register with the cluster of computers.

1   18.   A method for detecting and resolving a partition condition in a cluster of

2   computers in a networked environment, the method comprising:

3   creating a scratch pad area accessible by the cluster of computers;

4   dividing the scratch pad into a plurality of slots, each slot associated with a

5   plurality of nodes within the cluster of computers;

6   recording in the plurality of slots, a generation number and a list of known nodes

7   by each one of the plurality of nodes, wherein an identifier is written in the list for each

8   node that is known to a writing node and wherein the generation number and the list of

9   known nodes is recorded when a change of membership occurs in the cluster of

10 computers;

11   comparing each slot of the plurality of slots to ensure the generation number and

12 the list of known nodes matches in each slot of the plurality of slots; and

13   creating a list of surviving nodes by listing a first set of nodes determined by

14 comparing each slot of the plurality of slots;

15   re-allocating appropriate resources to each of the surviving nodes; and.

16   shutting down each node not on the list of surviving nodes by requiring each node

17 not on the list of surviving nodes to write a special message in a respective slot for that

18 node and then shut down immediately

1   19.   The method of claim 18 wherein the comparing each slot includes finding a list

2   with a master node to create the list of surviving nodes and shutting down each node not

3   on the list with the master node.

1   20.   The method of claim 18 wherein the comparing each slot includes finding a list

2   with a lowest node rank to create the list of surviving nodes and shutting down each node

3   not on the list with the lowest node rank.

1   21.   The method of claim 18 wherein the comparing each slot includes finding a list

2   with a largest node to create the list of surviving nodes and shutting down each node not

3   on the list with the largest node.

1   22.   The method of claim 18 wherein the comparing each slot includes finding a list

2   with a maximum number of nodes to create the list of surviving nodes and shutting down

3   each node not on the list with the maximum number of nodes.

1   23.   The method of claim 18 further including sending the list of surviving nodes to

2   the each node on the list of surviving nodes along with a new generation number.

1   24.   The method of claim 23 further including requiring each node not on the list of

2   surviving nodes to re-register with the cluster of computers.